Railroads and American Economic Growth: A "Market Access" Approach*

Dave Donaldson
MIT and NBER

Richard Hornbeck
Harvard and NBER

Draft Version: March 2012
Preliminary and Incomplete

Abstract

This paper examines the historical impact of railroads on the American economy. Expansion of the railroad network and decreased trade costs may affect all counties directly or indirectly, an econometric challenge in many empirical settings. However, the total impact on each county can be summarized by changes in that county's "market access," a reduced-form expression derived from general equilibrium trade theory. We measure counties' market access by constructing a network database of railroads and waterways and calculating lowest-cost county-to-county freight routes. As the railroad network expanded from 1870 to 1890, changes in market access are capitalized in agricultural land values with an estimated elasticity of 1.5. Removing all railroads in 1890 would decrease the total value of US agricultural land by 73% and GNP by 6.3%, more than double social saving estimates (Fogel 1964). Fogel's proposed Midwestern canals would mitigate only 8% of losses from removing railroads.

*For helpful comments and suggestions, we thank seminar participants at Berkeley, Harvard, Stanford, UC-Davis, and UC-Irvine. We are grateful to Jeremy Atack and coauthors for sharing their data. Irene Chen, Andrew Das Sarma, Manning Ding, Jan Kozak, Meredith McPhail, and Rui Wang provided excellent research assistance.

During the 19th century, railroads spread throughout a growing United States as the economy rose to global prominence. Railroads became the dominant form of freight transportation and areas around railroad lines prospered. The early historical literature often presumed that railroads were indispensable to the United States' economy or, at least, very influential for economic growth. Understanding the development of the American economy is shaped by an understanding of the impact of railroads and, more generally, the impact of market integration.

In *Railroads and American Economic Growth*, Fogel (1964) transformed the academic literature by using a social saving methodology to focus attention on counterfactuals: in the absence of railroads, freight transportation by rivers and canals would have been only moderately more expensive along most common routes. Small differences in freight rates can cause some areas to thrive relative to others, but the aggregate economic impact may be small. This social saving methodology has been widely applied to transportation improvements and other technological innovations, though many scholars have discussed both practical and theoretical limitations of the approach (see, e.g., Lebergott 1966; Nerlove 1966; McClelland 1968; David 1969; White 1976; Fogel 1979).[1]

There is an appeal to a methodology that estimates directly the impacts of railroads, using increasingly-available county-level data and digitized railroad maps. Recent work has compared counties that received railroads to counties that did not (Haines and Margo 2008; Atack and Margo 2010; Atack et al. 2010; Atack et al. 2011), and similar methods have been used to estimate impacts of railroads in modern China (Banerjee et al. 2010) or highways in the United States (Baum-Snow 2007; Michaels 2008). These studies estimate relative impacts of transportation improvements; for example, due to displacement and complementarities, areas without railroads and areas with previous railroads are also affected when railroads are extended to new areas.

This paper develops an alternative methodology for estimating the aggregate impact of railroads on the American economy. Expansion of the railroad network affects each county's "market access," a reduced-form expression derived from general equilibrium trade theory. County market access increases when it becomes cheaper to trade with another county, particularly when that other county has a larger population and higher trade costs with other counties. Rather than focus on changes in local railroad density, which may also be endogenous to local economic conditions, changes in market access summarize the total direct and indirect impacts on each county from changes in the national railroad network.[2]

[1] One alternative approach is to create a computational general equilibrium model, with the explicit inclusion of multiple regions separated by a transportation technology (e.g., Williamson 1974; Herrendorf et al. 2009).

[2] In related work using a similar model, Redding and Sturm (2007) estimate the impact on population

We measure counties' market access by constructing a network database of railroads and waterways and calculating lowest-cost county-to-county freight routes.

As the railroad network expanded from 1870 to 1890, changes in market access are capitalized in agricultural land values with an estimated elasticity of 1.5. An advantage of estimating the impact of market access, compared to the impact of local railroad density, is that counties' market access is influenced by changes elsewhere in the railroad network; indeed, this estimated elasticity is robust to controlling for changes in counties' own railroad track. Further, the estimated elasticity is robust to instrumenting for changes in market access with counties' initial market access through waterways only.

The paper then compares the observed impact of railroads in 1890 against two counterfactual scenarios considered by Fogel (1964). In the first scenario, removing all railroads is predicted to lower the total value of US agricultural land by 73% and generate an annual agricultural economic loss of 6.3% of GNP, roughly double social saving estimates by Fogel (1964). In the second scenario, replacing railroads with Fogel's proposed Midwestern canals is predicted to mitigate only 8% of the loss from removing railroads. Fogel's proposed canals are a poor substitute for the railroad network, which provides more-direct routes to more markets and larger decreases in distances of expensive wagon transportation.

Revisiting the historical impact of railroads on the American economy suggests a larger role for railroads and market integration in economic development. In the United States and around the world, considerable resources might be directed toward improving transportation infrastructure and market integration.

This paper also addresses the general methodological challenge of estimating aggregate treatment effects in empirical settings with substantial treatment spillover effects. Railroads affect all areas to some degree through interlinked trade networks. If railroads' impacts were locally confined, then the unit of analysis might be aggregated (e.g., Miguel and Kremer 2004). As in many empirical settings, however, sufficient aggregation is empirically intractable. The proposed solution uses economic theory to characterize how much railroads change each area's market access; once the intensity of treatment is defined to reflect both direct and indirect impacts, relative empirical comparisons estimate the aggregate treatment effect of railroads.[3]

from changes in market access following the division and reunification of Germany, and Hanson (2005) studies the correlation between US county-level wages and county-level market access from 1970 to 1990. Donaldson (2010) uses a different model to calculate overall gains from trade, estimating the impact of railroads on price dispersion, trade flows, and real income in colonial India.

[3]In the absence of an economic model, the spatial econometrics literature provides estimators for when treatment spillovers are a known function of geographic or economic distance (Anselin 1988). Estimation of aggregate treatment effects requires a cardinal ranking of how much areas (or people) are exposed to the treatment; an ordinal ranking is insufficient.

I Historical Background

During the 19th century, there was a dramatic shift westward in the geographic pattern of agricultural output. Increased regional trade surpluses and deficits in agricultural goods reflect the exploitation of comparative advantage. Fogel (1964) calculates the social saving from railroads in the agricultural sector in 1890, dividing the analysis between interregional trade and intraregional trade.

For interregional trade, from 9 primary markets in the Midwest to 90 secondary markets in the East and South, freight costs were only moderately cheaper with the availability of railroads than when using only natural waterways and canals. Multipling the difference in freight costs by the quantity of transported agricultural goods, Fogel calculates the annual interregional social saving to be $73 million or 0.6% of GNP. The total cost of agricultural interregional shipments would have nearly doubled, yet the aggregate economic impact is fairly small.

For intraregional trade, from farms to primary markets, railroads' impact was mainly in reducing distances of expensive wagon transportation. In the absence of railroads, farms would have incurred substantially higher costs in transporting goods by wagon to the nearest waterway to be shipped to the nearest primary market. Transportation costs by wagon would begin to exceed the average value of agricultural goods for areas more than 40 miles from a waterway; as an alternative, economic losses in this infeasible region are bounded by the value of agricultural land. Summing the annual return on the value of agricultural land in the infeasible region ($154 million) and increased transportation costs in the feasible region ($94 million), Fogel calculates the annual intraregional social saving to be $248 million or 2.1% of GNP.

With the aid of modern computers, finer distance buffers and county agricultural land values can be used to calculate the railroads' intraregional impact in both the infeasible region and the feasible region. Extending Fogel's use of the theory of rents, we assume that agricultural land loses a proportion of its value that declines linearly between 40 miles and 0 miles from the nearest waterway or railroad.[4] We multiply the value of agricultural land in each county by the implied decline in land value without railroads, based on the measured share of each county beyond 40 miles of a waterway, within 40 to 35 miles of a waterway, ..., and within 5 to 0 miles of a waterway.[5] Using Fogel's preferred rate of return, we calculate

[4]In practice, we assume that agricultural land loses 100% of its value beyond 40 miles, 93.75% of its value between 40 and 35 miles, 81.25% of its value between 35 and 30 miles, and so forth until losing 6.25% of its value between 5 and 0 miles.

[5]To avoid overstating the impact of railroads, we modify this calculation to also reflect counties' imperfect access to railroads. We use the county land share within each 5-mile distance buffer of a waterway, subtracting the county land share within that buffer of a waterway or railroad. To calculate percent declines off the

the annual intraregional impact of railroads to be $480 million or 4.0% of GNP.

Fogel's social saving methodology requires various assumptions and aggregations that might overstate or understate railroads' impact, yet Fogel's generally-maintained assumption is that railroads' impact is overstated because economic activity could be relocated in the absence of railroads. Further, canals might have been constructed through portions of the Midwest. The largest component of total social saving is the value of agricultural land more than 40 miles from a waterway, and a hypothetical network of feasible Midwestern canals would bring 70% of the lost land value within 40 miles of a waterway.[6]

Figure 1 illustrates the potential influence of additional Midwestern canals. In Panel A, areas shaded light gray are within 40 miles of a railroad in 1890 but not within 40 miles of a waterway (shaded black). In Panel B, areas shaded dark gray are further than 40 miles from a waterway but within 40 miles of Fogel's proposed canals. Nearly all Midwestern areas within 40 miles of railroads and not waterways are within 40 miles of the proposed canals; nationally, the same is true for most agricultural land by value.

Panels C and D replace the 40-mile buffers with 10-mile buffers, which reflect the average wagon haul from a farm to a rail shipping point in 1890. In contrast to Panels A and B, much more land is near railroads and not near waterways (Panel C) and the proposed canals are less effective in bringing farmland within 10 miles (Panel D). The high density of railroad construction is particularly effective in substituting for expensive wagon transportation. Railroads also provide more-direct routes to more destinations, compared to waterways and proposed canals.

Our empirical analysis maintains Fogel's focus on the agricultural sector. Railroads' impact on non-agricultural sectors is expected to be much smaller, as non-agricultural freight was geographically concentrated in areas with low transportation costs along waterways. In future versions of the paper, we plan to attempt to quantify the economic gains for consumers and non-agricultural sectors.

Our empirical analysis builds on Fogel's intuition that lower trade costs and improved market access are reflected in agricultural land values. We consider how expansion of the railroad network affects counties' market access and the associated impacts on agricultural land values. We then calculate the implied economic impact of decreased market access if railroads were eliminated or replaced with Fogel's proposed canals. There has been extensive debate regarding Fogel's social saving methodology; while we do not relitigate these issues,

correct base, we adjust observed county agricultural land values to reflect their implied value if not for distance to a waterway or railroad.

[6]Replicating our analysis of distance buffers and including the additional Midwestern canals, we calculate the annual intraregional impact of railroads to be $128 million or 1.1% of GNP. For this computational approach, Fogel's proposed canals mitigate 73% of the intraregional losses from removing railroads.

we discuss some topics as they relate to our alternative methodological approach.

II Data Construction

This paper uses a new dataset on predicted county-to-county freight transportation costs, calculated using a newly-constructed geographic information system (GIS) network database. This network database is similar to "Google Maps for the 19th century," incorporating the location of railroads in 1870 and 1890 and the location of canals and natural waterways. The end result is not the best prediction of actual freight costs, which might draw on published freight costs, but an econometrically-useful proxy for changes in freight costs due to expansion of the railroad network.

II.A Transportation Cost Parameters

Following Fogel's choice of transportation cost parameters, we set railroad rates equal to 0.63 cents per ton-mile and waterway rates equal to 0.49 cents per ton-mile.[7] Transshipment costs 50 cents per ton, incurred whenever transferring goods to/from a railroad car, river boat, canal barge, or ocean liner.[8] Wagon transportation costs 23.1 cents per ton-mile, defined as the straight line distance between two points.[9] Because railroad rates are similar to waterway rates and much smaller than wagon rates and transshipment costs, the most important aspects of network database construction concern the required distances of wagon transportation and number of transshipments.

We assume that freight rates per ton-mile are constant throughout the country. This assumption partly reflects data limitations, but local freight rates are likely endogenous to local economic activity and market power. Thus, for estimation purposes, there are advantages to using Fogel's average national rates. We hold rates fixed in 1890 and 1870, so measured changes in trade costs and market access are determined by variation in the location of transportation routes rather than prices.

II.B Transportation Network Database

Creation of the network database begins with digitized maps of constructed railroads in 1870 and 1890. We are grateful to Jeremy Atack and co-authors for providing these initial GIS

[7]Rates reflect an output-weighted average of rates for transporting grain and meat. Waterway rates include insurance charges for lost cargo (0.025 cents), inventory and storage costs for slower transport and non-navigable winter months (0.194 cents), and the social cost of public waterway investment (0.073 cents).

[8]Fogel considers transshipment charges as a sub-category of water rates, but our modeling of transshipment points allows for a unified treatment of Fogel's interregional and intraregional scenarios. Fogel's sources record higher railroad freight costs per ton-mile for shorter routes, but approximately reflect a 100 cent fixed fee and a 0.63 cent fee per mile.

[9]This rate reflects a cost of 16.5 cents per mile traveled and Fogel's adjustment factor of 1.4 between the shortest straight line distance and miles traveled.

railroad files.[10] These railroad files were originally created to define mileage of railroad track by county and year; by contrast, for our purposes, railroad lines are modified to ensure that GIS software recognizes travel is possible through the railroad network.[11]

The second step adds the time-invariant locations of canals, navigable rivers, and other natural waterways. We use Fogel's definition of navigable rivers, which are enhanced to follow natural river bends. For lakes and oceans, we saturate their area with "rivers" that allow for a large number of possible routes. Transshipment costs are incurred whenever freight is transferred between the four transportation methods: railroad, canal, river, and lake or ocean.[12]

The third step connects individual counties to the network of railroads and waterways. We measure average travel costs between counties by calculating the travel cost between county centroids. County centroids must be connected to the network of railroads and waterways; otherwise, lowest-cost travel calculations assume that freight travels freely to the closest railroad or waterway. We create wagon routes between each county centroid and each nearby type of transportation route in each relevant direction. Because the network database only recognizes lines, we also create direct wagon routes from every county centroid to every other county centroid within 300km.[13]

The fourth step refines centroid-to-network connections due to the importance of wagon distances to overall freight costs. For example, when a railroad runs through a county, the centroid's nearest distance to a railroad does not reflect the average distance from county points to a railroad.[14] We create 200 random points within each county, calculate the distance from each point to the nearest railroad, and take the average of these nearest distances. We then adjust the cost of travel along each centroid connection to within-county railroads

[10]First, year-specific railroad atlas scans from the Library of Congress are "georeferenced" to US county borders. Second, railroad lines are hand-traced in GIS software to create a digital map of railroad line locations.

[11]We use GIS topology tools to ensure exact connections between all railroad line segments. Hand-traced railroad lines often contain small internal gaps and, by default, intersecting lines are equivalent to a highway overpass without a connection. These modifications to the railroad network have little effect on total railroad track mileage by county and year. To minimize measurement error in changes, we created a final 1890 railroad file and modified that file to create a version for 1870 that omits lines constructed between 1870 and 1890.

[12]Overlapping railroads and waterways do not connect by default; instead, we create connections among railroads and waterways to allow for fixed transshipment costs. The need to include transshipment costs is the main reason why it is not possible to model the network using a raster, assigning travel costs to each map pixel.

[13]The direct wagon routes are restricted to be over land, but there is no adjustment for mountains or other terrain; in practice, the long-distance wagon routes are very costly and unused for most of the sample counties.

[14]Fogel recognized the importance of measuring this within-county distance and his ideal solution was to break each county into small grids and take the average of nearest distances from each grid to a railroad. Due to technical limitations, Fogel approximated this average distance using one-third of the distance from the farthest point in a county to a railroad.

to reflect that county's average travel cost to a railroad. We then repeat this procedure for centroid connections to navigable rivers and canals. This refinement to the network database allows the empirical analysis to exploit more-precise variation on the intensive margin of county access to railroads and waterways as the density of the railroad network increases from 1870 to 1890.

Figure 2 shows part of the created network database. Panel A shows all natural waterways, including the navigable rivers and routes across lakes and oceans. Panel B adds the canal network, which is highly complementary with natural waterways. Panel C adds railroads constructed in 1870, and then Panel D adds railroads constructed between 1870 and 1890. Early railroads were complementary with the waterway network; by 1870 and especially by 1890, the railroad network is more of a substitute for the waterway network.

II.C Limitations of the Network Database

There are several limitations of the constructed network database. Freight rates are not allowed to vary by direction, reflecting either river currents or back-haul trade relationships. Freight rates also do not vary with local terrain or market power in the transportation sector. Predicted railroad routes do not reflect railroad line ownership or gauges, so actual routes may be more circuitous or delayed. There are no congestion effects or scale economies in processing goods, so the network database is not suited to predicting transportation hubs. These are important limitations, though even large percent differences in railroad and waterway costs will have relatively little impact on overall transportation costs: 10 miles of wagon transportation are roughly equivalent to 375 – 475 miles of railroad or waterway transportation. Further, the empirical analysis can explore the results' robustness to alternative cost parameters.

Inaccuracies in the network database that overstate the impact of railroads on trade costs will attenuate the estimated impact of market access on land values. When evaluting the economic impact of removing railroads, however, the decline in market access will be overstated and counteract the downward-biased elasticity. Thus, the overall predicted economic losses from removing railroads should be less-affected by mismeasurement than the intermediate estimates of how the railroad network affects market access or how market access affects land values.

II.D Transportation Route Cost Calculations

The complete network database allows the calculation of lowest-cost routes between each pair of counties, i.e., 5 million calculations. The lowest-cost routes are calculated under four scenarios: (1) the waterway and railroad network in 1870, (2) the waterway and railroad network in 1890, (3) the waterway network only, and (4) the waterway network with additional

Midwestern canals. The created data are not our best predictions of actual freight routes and costs, but an econometrically-useful proxy for differences over space and time due to the differences in the location of railroads and waterways.

II.E County-level Census Data

County-level data are drawn from the US Censuses of Agriculture and Population (Haines 2005). The main variables of interest are the value of agricultural land and the total population. We adjust data from 1870 to reflect 1890 county boundaries (Hornbeck 2010).

III A "Market Access" Approach to Valuing Railroads

The empirical analysis is guided by a model of trade among US counties that specifies how each county is affected by changes in the national matrix of county-to-county trade costs. The model contains thousands of counties, each with interacting goods markets and factor markets, that generate positive and negative spillovers on other counties. Under a set of assumptions that are standard among modern quantitative trade models, all direct and indirect impacts of changing trade costs are reflected, in equilibrium, in changes to a county's "market access."[15]

The model implies a simple log-linear relationship between county agricultural land values and county market access, appropriately defined. While the model requires particular assumptions to arrive at this parsimonious solution to the challenges posed by general equilibrium spatial spillovers, the predicted relationship also has an atheoretical appeal in capturing the impact of railroads. County market access increases when it becomes cheaper to trade with another county, particularly when that other county has a larger population and higher trade costs with other counties. Guided by the model, we regress agricultural land value on market access and a set of controls.

III.A A Model of Trade Among US Counties

The economy consists of many trading counties, each indexed by o if the origin of a trade and by d if the destination. Some counties are rural ($o \in \mathcal{R}$) and some are urban ($o \in \mathcal{U}$).

Tastes: Economic agents have Cobb-Douglas preferences over two commodity groups: agricultural goods (weight γ) and manufactured goods (weight $1 - \gamma$). The commodity groups are composed of many differentiated goods varieties (indexed by j), and tastes over these varieties take a CES form (with elasticities σ_A and σ_M).[16] An agent living in county

[15] These modeling assumptions are used in the fields of international trade and economic geography, and reflect recent best-practice to gain traction in general equilibrium spatial settings with many regions and commodities.

[16] Note that the elasticities of substitution within each commodity group are not restricted, i.e., these could be high if agricultural varieties are relatively similar. Anderson et al. (1992) provide an attractive microfoundation for aggregate-level CES preferences: if individual agents desire only one variety of the

o and receiving income Y_o experiences indirect utility:

$$V(\mathbf{P_o}, Y_o) = \frac{Y_o}{(P_o^A)^\gamma (P_o^M)^{1-\gamma}}, \tag{1}$$

where P_o^A and P_o^M are standard CES price indices.[17]

Agricultural technology: For simplicity, we assume a stark form of specialization: rural counties produce only agricultural goods and urban counties produce only manufacturing goods. Rural counties use a Cobb-Douglas technology to transform land and labor into agricultural goods varieties. We assume that labor is mobile across space and across sectors, i.e., that farmers have an endogenous outside option to work in other counties or in manufacturing. The marginal cost of producing agricultural goods of variety j in county o is:

$$MC_o^A(j) = \frac{(q_o)^\alpha (w_o)^{1-\alpha}}{z_o^A(j)} \quad \text{if } o \in \mathcal{R}, \tag{2}$$

where w_o is the wage rate, q_o is the agricultural land rental rate, and $z_o^A(j)$ is a Hicks-neutral productivity shifter that is exogenous and local to county o. We follow Eaton and Kortum (2002) in modeling these productivity shifters by assuming that each county draws its productivity level, for any given variety j, from a Frechet (or Type II extreme value) distribution: $F_o(z) = 1 - \exp(-T_o^A z^{-\theta_A})$.[18] Finally, we assume perfect competition among producers, as is natural for the case of relatively homogeneous goods such as agricultural commodities.

Manufacturing technology: Urban counties use a Cobb-Douglas technology to transform capital and labor into manufactured goods varieties, whose marginal cost of production is:

$$MC_o^M(j) = \frac{(w_o)^\beta (r_o)^{1-\beta}}{z_o^M(j)} \quad \text{if } o \in \mathcal{U}, \tag{3}$$

agricultural good (their "ideal variety") and agents' ideal varieties are distributed in a Logit fashion, then aggregate consumption data from a population of many such agents behaves as though all agents have CES preferences over all varieties (where in such an interpretation σ indexes the inverse of the dispersion of agents' ideal varieties).

[17]That is, $P_o^A = \left[\int_0^{n^A} (p_c^A(j))^{1-\sigma_A} dj\right]^{1/(1-\sigma_A)}$ and $P_o^M = \left[\int_0^{n^M} (p_o^M(j))^{1-\sigma_M} dj\right]^{1/(1-\sigma_M)}$. Variables n^A and n^M are the number of varieties available to consumers; as we describe below, the number of varieties is fixed in agriculture and the number is endogenous in the manufacturing sector and adjusts to allow free entry.

[18]This distribution captures how productivity differences across counties give incentives to specialize and trade. The parameter T_o^A captures county-specific (log) mean productivity, which corresponds to each county's level of absolute advantage. The parameter θ^A captures, inversely, the (log) standard deviation of productivity, which corresponds to the scope for comparative advantage. A low θ_A means county productivity draws are dispersed, creating large incentives to trade on the basis of productivity differences.

where w_o is the same wage rate, r_o is the capital rental rate, and $z_o^M(j)$ is a Hicks-neutral productivity shifter that is exogenous and local to county o. The urban manufacturing sector pays both this marginal cost and a fixed cost of production, f_o, paid in the same proportion of factors as marginal costs. We follow Melitz (2003) in assuming a set of potential producers, each with a productivity shifter $z_o^M(j)$ that is drawn, prior to entry, from a distribution with CDF $G_o(z)$. We also follow Chaney (2008), among others, in modeling the distribution $G_o(z)$ as a Pareto distribution: $G_o(z) = 1 - (z/T_o^M)^{-\theta_M}$.[19] The parameter T_o^M captures the location of this distribution and θ_M captures its dispersion. Following Melitz (2003), firms compete monopolistically with free entry such that all firms' expected profits are zero.

Trade costs: Trading goods is costly. Remote rural locations pay high prices for imported goods and receive low prices for goods they produce, as this is the only way locations can be competitive in distant markets. We model trade costs using a simple and standard "iceberg" formulation. When a variety is made in county o and sold locally in county o, its price is $p_{oo}^k(j)$ in sector k; but when this same variety is made in county o and shipped to county d, it will sell for $p_{od}^k(j) = \tau_{od}^k p_{oo}^k(j)$.[20] A proportional trade cost τ_{od}^k is applied to each unit of the variety shipped. Trade is costly because $\tau_{od}^k \geq 1$.

Factor mobility: Workers are free to move around the country, at least over a period of many years. Worker utility, as defined in equation (1), is equalized across counties in equilibrium and wages satisfy:

$$w_o = \bar{U}(P_o^A)^\gamma (P_o^M)^{1-\gamma}, \quad (4)$$

where $\bar{U}$ is the endogenous level of utility obtained by workers in each county. Land is fixed by county and capital is perfectly mobile.

III.B Solving the Model

Trade flows and the gravity equation: First, we solve for agricultural goods' trade flows from each origin county o to each other destination county d. Due to perfect competition, the marginal costs of producing each variety is equal to its price. We assume that consumers only buy a variety from the cheapest available location. Substituting marginal costs from each supply location o (equation 2) into the demand for agricultural varieties in county d,

[19] This is a good match for the empirical distribution of firm sizes in most samples (e.g., Axtell 2001).

[20] Monopolistically competitive producers in the manufacturing sector have an incentive to segment markets but, with CES preferences and many producers, the same mark-up over marginal cost will be charged by each producer in each market.

Eaton and Kortum (2002) derive the value of total exports from o to d:[21]

$$X_{od}^{A} = \chi_A T_o^A (q_o^{\alpha} w_o^{1-\alpha})^{-\theta_A} (\tau_{od}^A)^{-\theta_A} (CMA_d^A)^{-1} Y_d, \tag{5}$$

$$\text{where} \quad CMA_d^A \equiv \sum_{i \in \mathcal{R}} T_i^A (q_i^{\alpha} w_i^{1-\alpha})^{-\theta_A} (\tau_{id}^A)^{-\theta_A}. \tag{6}$$

CMA_d^A refers to "consumer market access" for agricultural goods in destination market d. Consumer market access is a weighted sum of productivity-adjusted costs of production in each market i that could supply market d, with weights declining in the cost of trading from i to d (i.e., τ_{id}^A). From equation (5), county o sends more goods to county d if county o is relatively productive (high T_o^A) or relatively low cost (low q_o or w_o). County o also sends more goods to county d if county d has high total income (high Y_d) or low overall consumer market access (low CMA_d^A), meaning that county o faces fewer competitors when selling to market d.

A similar expression can be derived for trade flows in the manufacturing sector. While the manufacturing sector's solution is complicated by the presence of free entry, Chaney (2008) shows that trade flows again take a gravity equation form:

$$X_{od}^{M} = \chi_M T_o^M n_o^M (w_o^{\beta} r_o^{1-\beta})^{-\theta_M} (\tau_{od}^M)^{-\theta_M} (CMA_d^M)^{-1} Y_d, \tag{7}$$

$$\text{where} \quad CMA_d^M \equiv \sum_{i \in \mathcal{U}} T_i^M n_i^M (w_i^{\beta} r_i^{1-\beta})^{-\theta_M} (\tau_{id}^M)^{-\theta_M}. \tag{8}$$

The only difference from equation (5) is that manufacturing exports from county o also depend on the endogenous number of varieties available in each location (i.e., n_o^M).

Equations (5) and (7) are known as gravity equations, and they govern trade flows in this model. Deriving a gravity equation is dependent on the particular functional form assumptions for tastes and technology; indeed, it is difficult to imagine alternative assumptions that give rise to gravity equations in these settings (Arkolakis et al. 2011). The gravity equation is appealing because it dramatically simplifies a complex general equilibrium problem of spatial competition. A more appealing feature of the gravity equation is that it appears to fit trade-flow data in many contexts (e.g., Anderson and van Wincoop 2003, 2004; Combes et al. 2008). Thus, it is difficult to imagine an exercise that models trade among counties in the United States that does not predict trade flows will take the "gravity" form.

Land rental rate: While trade flows between 19th century US counties are unobserved, the gravity equation implies tractable expressions for outcome variables that are observed and important for our purposes, such as the agricultural land rental rate (q_o).

[21] The term χ_A is a constant that is comprised of the parameters γ, θ_A, and σ_A.

Under the assumption of Cobb-Douglas agricultural technology, agricultural land is paid a fixed share α of total output Y_o^A, so $q_o L_o = \alpha Y_o^A$ where L_o is the (exogenous) quantity of agricultural land in county o. Goods markets clear, so all produced goods are bought ($Y_o^A = \sum_d X_{od}^A$). Thus, equation (5) implies:

$$q_o^{1+\alpha\theta_A} = \alpha \left(\frac{T_o^A}{L_o} \right) w_o^{-\theta_A(1-\alpha)} FMA_o^A. \tag{9}$$

FMA_o^A refers to "firm market access" for agricultural goods from origin o and is defined as:

$$FMA_o^A \equiv \gamma \sum_{d \in \mathcal{R},\mathcal{U}} (\tau_{od}^A)^{-\theta_A} (CMA_d^A)^{-1} Y_d. \tag{10}$$

Firm market access (FMA_o^A) is a sum of terms over all destination counties d to which the rural county o tries to sell its agricultural goods. These terms include the size of the destination market (given by total income, Y_d) and the competitiveness of the destination market (given by its CMA_d^A term). All terms are inversely weighted by the cost of trading with each distant market ($(\tau_{od}^A)^{-\theta_A}$).

Substituting equation (4) into equation (9) to eliminate the wage, we obtain an expression for the land rental rate in county o:

$$\begin{aligned} \ln q_o = \kappa &+ \left(\frac{1}{1+\alpha\theta_A} \right) \ln \left(\frac{T_o^A}{L_o} \right) + \left(\frac{\gamma(1-\alpha)}{1+\alpha\theta_A} \right) \ln(CMA_o^A) \\ &+ \left(\frac{(1-\gamma)(1-\alpha)}{1+\alpha\theta_A} \right) \ln(CMA_o^M) + \left(\frac{1}{1+\alpha\theta_A} \right) \ln(FMA_o^A), \end{aligned} \tag{11}$$

where κ is a constant that subsumes model parameters and $\bar{U}$.

III.C Using the Model to Inform Empirical Work

Equation (11) provides a useful guide for the empirical analysis. Equilibrium rural land rental rates (q_o) are log-linear in just three endogenous economic variables: (1) firm market access for agricultural goods that the rural area produces (FMA_o^A), (2) consumer market access to agricultural goods (CMA_o^A), and (3) consumer market access to manufactured goods (CMA_o^M). Immobile land in county o will be more valuable if county o has: cheap access to large uncompetitive markets (high FMA_o^A) or cheap access to agricultural goods (high CMA_o^A) and manufactured goods (high CMA_o^M) that workers value.

Equation (11) has two key implications for estimating the aggregate economic impact of railroads. First, all economic forces that make goods and factor markets interdependent across counties are represented by three slightly different concepts of "market access." Thus, all direct and indirect impacts of railroads are captured by analyzing changes in this particular concept of market access. For example, county A receiving a railroad line would affect other counties: those that can now trade with county A, those that had been trading with county A, those that had traded with county A's previous trade partners, those that had traded with county A's new trade partners, and so on. Even if access to railroads is randomly assigned, "control" counties are affected and a regression of land rents on railroad access will produce biased estimates of railroads' impact. A regression of land rents on market access, however, will be free of this bias, in the context of our model, because all counties' market access will adjust to changes in the railroad network.

The second key implication of equation (11) is that a county's market access can increase or decrease due to changes in the railroad network far beyond that county's borders. Thus, the empirical estimation is not identified only from particular counties gaining railroad access, which might otherwise be correlated with land rental rates. This prediction of the model suggests several control variables or instrumental variable approaches that might purge the empirical estimates of endogeneity bias arising from railroad placement decisions.

III.D From Theory to an Empirical Specification

While equation (11) provides a useful guide for the empirical analysis, three obstacles prevent its direct empirical implementation.

First, the Census of Agriculture does not report land rental rates; instead, the Census reports the value of agricultural land and buildings.[22] We use land values as a proxy for land rental rates.[23] While land values may reflect expected changes in market access, this concern may be lessened by controlling for changes in local railroads that are more-easily anticipated. If agents partly anticipate changes in market access, it will attenuate the estimated elasticity.

Second, county agricultural productivity (T_o^A) is not directly observed. We assume that county productivity is constant between 1870 and 1890, or that changes in county productivity are orthogonal to changes in market access. In practice, we allow for changes in county productivity to vary by state and by cubic polynomials in county latitude and longitude.[24]

Third, and most importantly, the three market access terms (FMA_o^A, CMA_o^A, CMA_o^M) are also not directly observed. Each market access term contains variables that are unob-

[22]In years when the Census reports separate values for land and buildings, land value is the largest share of the combined measure. If agricultural buildings are fixed to the land, then changes in market access will also change the value of agricultural buildings.

[23]In particular, we assume that $V_o = q_o/r$, where V_o is the land value and r is a fixed interest rate. In practice, the interest rate can vary by county, state-year, or with any of the control variables in the empirical specification.

[24]Because county productivity (T_o^A) enters log-linearly in equation (11), we control for this term using county fixed effects, state-by-year fixed effects, and year-interacted cubic polynomials in county latitude and longitude.

served: the wage w_d, the capital rental rate r_d, and the destination technologies T_d^A and T_d^M. These three market access terms can be calibrated using population data, as done by Redding and Sturm (2007), though for now we implement a less-restricted approach.

We use one simplified market access term (MA_o) to proxy for the three similar unobserved market access terms from the model (FMA_o^A, CMA_o^A, CMA_o^M).[25] Similar to the FMA and CMA terms, this one measure of county "market access" is a sum over other counties' market sizes (X_{dt}), adjusted for the cost of trading with each other county ($\tau_{odt}^{-\theta}$) and adjusted for each other county's access to other markets ($\sum_{i \neq d,o} \tau_{dit}^{-\theta} X_{it}$):

$$MA_{ot} = \sum_{d \neq o} \frac{(\tau_{odt})^{-\theta} X_{dt}}{\sum_{i \neq d,o} (\tau_{dit})^{-\theta} X_{it}}. \tag{12}$$

We use county population as a measure of market size. We use the GIS-calculated county-to-county lowest-cost freight transportation routes as a measure of trade costs.[26]

County market access increases when it becomes cheaper to trade with another county, particularly when that other county has a larger population and higher trade costs with other counties. While this single market access term is more "ad hoc" than the three terms that follow directly from the model, MA_{ot} can be formally shown to approximate CMA_{ot}^A, CMA_{ot}^M and FMA_{ot}^A as the location of manufacturing and agriculture in the vicinity of county o becomes more and more similar. Indeed, it is natural that a farmer's market access is highly correlated with the farmer having lower trade costs with: (1) larger populations demanding the farmer's goods, (2) larger populations that have fewer other sources for acquiring the farmer's goods, (3) larger populations producing the farmer's demanded goods, and (4) larger populations that have fewer other destinations to sell the farmer's demanded goods. These four forces are embodied in equation (12).

Our measure of market access is partly inspired by an older concept of "market potential," based on the number and magnitude of markets available at low trade costs (Harris 1954). Harris's market potential term effectively equals $\sum_{d \neq o} (\tau_{odt})^{-1} X_{dt}$, where X_{dt} is a proxy for total demand at location d. We extend this formula to deflate market sizes by their competitiveness: a large market d is more appealing to exporters in county o if market d is served by fewer other counties and, thus, prices are high in market d; alternatively, access to

[25] Combining these three terms into one removes the model's distinction between urban and rural counties (and between agricultural and manufacturing sectors).

[26] We express the calculated trade costs in proportional terms using Fogel's average value of transported agricultural goods. The parameter θ_A is known as the "trade elasticity." Donaldson (2010) estimates a value of 3.8 in colonial Indian agriculture and Simonovska and Waugh (2011) estimate a value of 4.5 in modern OECD manufacturing. We follow Donaldson (2010) and use a value of $\theta_A = 3.8$; the empirical analysis can explore the results' robustness to alternative parameters.

a large market is more appealing to consumers in market o if market d supplies fewer other counties and, thus, sells goods for low prices. We also allow trade costs to diminish market sizes to the power of $-\theta_A$, rather than one, as typical estimates of θ_A are substantially greater than one.

Summarizing the above discussion, we regress agricultural land values in county o and year t on log market access (MA_{ot}), a county fixed effect (α_o), state-by-year fixed effects (α_{st}), and a cubic polynomial in county latitude and longitude interacted with year effects ($f(x_o, y_o)\gamma_t$):

$$\ln V_{ot} = \beta \ln(MA_{ot}) + \alpha_o + \alpha_{st} + f(x_o, y_o)\gamma_t + \varepsilon_{ot}. \tag{13}$$

The sample is a balanced panel of 2,161 counties with land value data in 1870 and 1890. Standard errors are clustered at the state level to adjust for heteroskedasticity and within-state correlation over time.

IV Empirical Results

IV.A Estimated Impact of Market Access on Land Values

Table 1 presents baseline results from estimating equation (13). Market access is estimated to have an economically large and statistically significant impact on land values, with an elasticity of 1.5 (Column 1).[27] Columns 2 – 6 allow the impact of market access to vary by region: the estimated impact of market access is largest in the Midwest and Plains, whereas changes in market access may be measured with more error in the Northeast and Far West.[28]

The main empirical concern is that expansion of the railroad network is endogenous, so increases in county market access may be otherwise correlated with increases in county land values. The baseline regressions do control for differential changes by state and county latitude and longitude, yet the empirical concern remains. There may be some exogenous influences on railroad placement, similar to those used in other papers on railroads and highways, though during our sample period there is a fairly high density of railroads outside

[27]When defining market sizes in equation (12) using county land values, rather than county population, the estimated elasticity is 1.44. Neither variable exactly matches total demand in other counties, as in the model, though the similarity in results is reassuring. The remainder of the results use population as a proxy for market size, as there are advantages to not including land values on both sides of the regression.

[28]Under the model approximation discussed above, i.e., as the location of manufacturing and agriculture in the vicinity of the county becomes more and more similar, the coefficient on the market access term should satisfy $\beta = \left(\frac{2-\alpha}{1+\alpha\theta}\right)$. For $\theta = 3.8$ and $\alpha = 0.19$, which reflects the land share in agricultural production (Caselli and Coleman 2001), the model approximation predicts $\beta = 1.05$. This formula only holds under an approximation to the theoretical model, and our empirical measure of market access is only an approximation as well, yet the predicted and estimated coefficients are fairly similar. The empirical estimates would not reject a coefficient of 1.05.

the Far West. We exploit two alternative identification strategies that take advantage of variation in market access that is not determined by local changes in the railroad network.

First, we focus on variation in county market access that is orthogonal to that county's railroads. Table 2, column 1 reports the baseline estimate for comparison. Conditional on the same controls from equation (13), whether a county has any railroad track is associated with a 0.359 log point increase in land values (column 2) and a 0.223 log point increase in county market access (column 3). However, the estimated impact of market access is remarkably robust to controlling for whether a county has any railroad track (column 4). These estimates are similar when also controlling for the length of railroad track by county (column 5).[29]

Thus, while county market access is correlated with railroad track, the estimated impact of market access is robust to exploiting variation in market access that is independent of that county's own railroad construction and associated endogeneity bias.[30] Further, once controlling for market access, the estimated impact of having any railroad track is no longer substantial or statistically significant. Railroad track is only estimated to increase land values to the extent that it increases county market access, consistent with our definition of market access being the correct functional form through which the railroad network affects counties.

Second, as an alternative identification strategy, we use the historical substitutability between railroads and waterways. In particular, expansion of the national railroad network is expected to have a larger impact on counties with worse market access through waterways. Regardless of how the railroad network actually changes from 1870 to 1890, counties with better market access through waterways in 1870 are expected to have a smaller increase in market access from 1870 to 1890. Using a restricted GIS network database with no railroads, we calculate county-to-county lowest-cost freight routes and measure counties' access to markets in 1870 through waterways only.

In the first-stage specification, Table 3 reports that counties with better "water market access" in 1870 experience a relative decline in market access from 1870 to 1890 (column 1). In the reduced-form specification, counties with better "water market access" in 1870 also experience a relative decline in agricultural land values (column 2). Instrumenting for the change in market access with counties' initial "water market access," market access

[29]The coefficient on market access is also robust to controlling for higher-order polynomials of railroad track length.

[30]Perhaps surprisingly, the estimated impact of market access is similar when using *only* variation in market access from that county's own railroad construction. While not the main purpose of this table, an instrumental variables estimate of the impact of market access is given by the ratio of the coefficient in column 2 (the reduced-form impact of railroad construction on land values) to the coefficient in column 3 (the first-stage impact of railroad construction on market access), which yields an IV estimate of 1.610.

is estimated to increase land values with an elasticity of 2.154. One concern is that a general trend toward convergence in both market access and agricultural land values may overstate the instrumented impact of market access; controlling for changes correlated with counties' initial value of agricultural land, the instrumented elasticity is estimated to be 1.271. Using this instrumental variables approach, the estimated elasticities are not statistically distinguisable from the baseline estimates in Table 1 (column 1).

IV.B Economic Impact of Removing Railroads in 1890

Using the estimated relationship between market access and land values, we calculate the predicted aggregate economic impact from removing all railroads in 1890. First, we use county-to-county lowest-cost freight routes in a restricted network database with no railroads to measure each county's counterfactual market access in 1890. Market access falls by 63%, on average, when all railroads are eliminated. By comparison, the average increase in market access from 1870 to 1890 was 50%; thus, the counterfactual decline in market access is not substantially "out of sample." The model also predicts a linear relationship between log land values and log market access; indeed, the impact of market access is estimated to be approximately linear.[31]

County agricultural land value is predicted to fall by 1.477 log points for every log point decline in market access (Table 1, column 1), and the implied percent decline in each county's land value is multiplied by each county's land value in 1890. Summing over all counties, removing all railroads in 1890 is predicted to decrease the total value of US agricultural land by 73% or $9.5 billion (1890 dollars). Using Fogel's preferred rate of return, the implied annual economic loss is $751 million or 6.3% of GNP.

IV.C Economic Impact of Replacing Railroads with Canals in 1890

We also calculate the economic impact of replacing all railroads with a hypothetical system of Midwestern canals. Fogel (1964) proposes a system of canals that might have substantially mitigated the economic losses from eliminating all railroads. When replacing all railroads with the proposed canal network, however, we calculate an annual economic loss of $692 million or 5.8% of GNP. The proposed canals are a poor substitute for the railroad network, mitigating only 8% of the losses from removing railroads. While canals reach within 40 miles of many Midwestern areas, the railroad network provides more-direct routes to more markets and larger decreases in distances of expensive wagon transportation. The proposed canals do generate $59 million in annual benefits, in the absence of railroads, which exceeds

[31] When including a fourth-degree polynomial in the log of market access, the coefficient on the linear term remains similar and the higher-order terms are jointly insignificant.

their estimated annual capital cost of $34 million.[32] However, the 1890 railroad network generated substantially larger economic benefits than the potential gains from feasible canal construction.

V Conclusion

(To Come)

[32]Paid freight rates would also partly cover canals' capital costs or even generate rents.

References

Anderson, S. P., A. de Palma, and J. Thisse (1992): *Discrete Choice Theory of Product Differentiation*. MIT Press, Cambridge.

Anderson, J. and E. van Wincoop (2003): "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, 93(1), pp. 170-192.

Anderson, J. and E. van Wincoop (2004): "Trade Costs," *Journal of Economic Literature* XLII(3), pp. 691-751.

Anselin, L. (1988): *Spatial Econometrics: Methods and Models*, Dordrecht: Kluwer Academic Publishers.

Arkolakis, C., A. Costinot, and A. Rodríguez-Clare (2011): "New Trade Models, Same Old Gains?," *American Economic Review*, forthcoming. http://econ-www.mit.edu/files/6445.

Atack, J., F. Bateman, M. Haines and R.A. Margo (2010): "Did Railroads Induce or Follow Economic Growth? Urbanization and Population Growth in the American Midwest, 1850-1860," *Social Science History*, 34(2), pp. 171-197.

Atack, J., M. Haines and R.A. Margo (2011): "Railroads and the Rise of the Factory: Evidence for the United States, 1850-1870," in P. Rhode, J. Rosenbloom, D. Weiman, eds. *Economic Evolution and Revolutions in Historical Time*. Palo Alto, CA: Stanford University Press.

Atack, J. and R.A. Margo (2010): "The Impact of Access to Rail Transportation on Agricultural Improvement: The American Midwest as a Test Case, 1850-1860," BU Mimeo, March. http://www.bu.edu/econ/files/2010/10/FinalDraftJTLUMarch-28-2010JournalVersion1.pdf

Axtell, R. (2001): "Zipf Distribution of U.S. Firm Sizes," *Science*, 293(5536), pp. 1818-1820.

Baum-Snow, N. (2007), "Did Highways Cause Suburbanization?," *The Quarterly Journal of Economics*, 122(2), pp. 775-805.

Caselli, F. and W. Coleman II (2001): "The U.S. Structural Transformation and Regional Convergence: A Reinterpretation," *The Journal of Political Economy*, 109(3), pp.584-616.

Chaney, T. (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, 98(4), pp. 1707-1721.

Combes, P-P., T. Mayer and J-F. Thisse (2008): *Economic Geography: the Integration of Regions and Nations*, Princeton: Princeton University Press.

David, P. (1969): “Transport Innovation and Economic Growth: Professor Fogel on and off the Rails,” *The Economic History Review*, 22(3), pp. 506-524.

Donaldson, D. (2010): “Railroads of the Raj: Estimating the Impact of Transportation Infrastructure,” unpublished manuscript, MIT. http://econ-www.mit.edu/files/6038.

Eaton, J. and S. Kortum (2002): “Technology, Geography and Trade,” *Econometrica*, 70(5), pp. 1741-1779.

Fogel, R. W. (1964): *Railroads and American Economic Growth: Essays in Economic History*, Johns Hopkins University Press, Baltimore.

Fogel, R. W. (1979): “Notes on the Social Saving Controversy,” *Journal of Economic History*, 39, pp. 1-54.

Haines, M. (2005): *Historical, Demographic, Economic, and Social Data: US, 1790-2000*, ICPSR.

Haines, M. and R.A. Margo (2008): “Railroads and Local Economic Development: The United States in the 1850s,” in J. Rosenbloom, ed., *Quantitative Economic History: The Good of Counting*, pp. 78-99. London: Routledge.

Hanson, Gordon H. (2005): "Market potential, increasing returns and geographic concentration," *Journal of International Economics*, 67(1), pp. 1-24.

Harris, C. (1954): “The Market as a Factor in the Localization of Industry in the United States,” *Annals of the Association of American Geographers*, 64, pp. 315-348.

Herrendorf, B. J.A. Schmitz Jr. and A.Teixeira (2009): "Transportation and Development: Insights from the U.S., 1840—1860," Federal Reserve Bank of Minneapolis Research Department Staff Report 425. http://minneapolisfed.org/research/sr/sr425.pdf

Hornbeck, R. (2010): “Barbed Wire: Property Rights and Agricultural Development,” *The Quarterly Journal of Economics*, 125, pp. 767-810.

Lebergott, S. (1966): "United States Transportation Advance and Externalities", *Journal of Economic History*, 26(4), pp 437-461.

McClelland, P. D. (1968): "Railroads, American Growth, and the New Economic History: A Critique", *Journal of Economic History*, 28(1), pp 102-123.

Melitz, M. (2003): “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity,” *Econometrica*, 71, pp. 1695-1725.

Michaels, G. (2008): "The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System," *The Review of Economics and Statistics*, 90(4), pp. 683-701.

Nerlove, M. (1966): "Railroads and American Economic Growth", *Journal of Economic History*, 26(1), pp 107-115.

Redding, S. and D. Sturm (2007): “The Cost of Remoteness: Evidence from German Division and Reunification,” *American Economic Review*, 98(5), pp. 1766-1797.

Simonovska, I. and M. Waugh (2011): “Elasticity of Trade: Estimates and Evidence,” Unpublished Manuscript, UC Davis and NYU. https://files.nyu.edu/mw134/public/uploads/56836/estimate_theta_paper.pdf.

White, C. M. (1976): "The Concept of Social Savings in Theory and Practice", *Economic History Review* 24, pp 82-100.

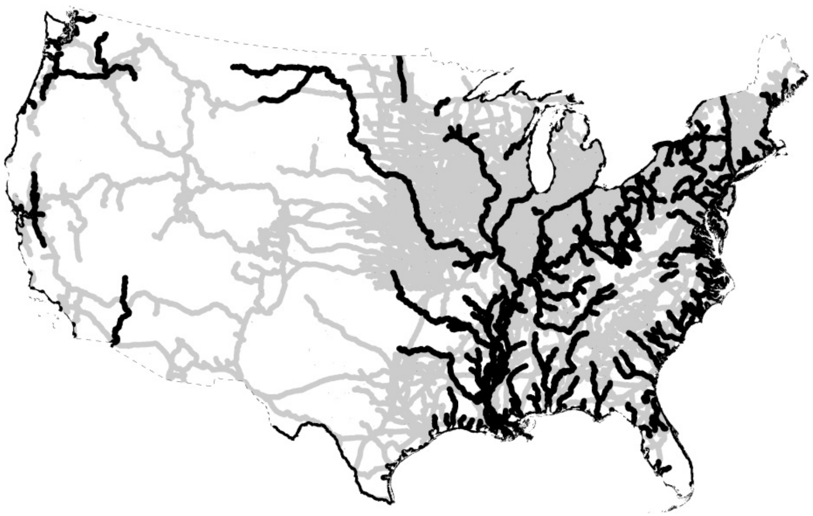
Williamson, J.G. (1974): *Late Nineteenth-Century American Development: a General Equilibrium History*, New York: Cambridge University Press.

Figure 1. Distance Buffers in 1890 around Waterways, Railroads, and Proposed Canals

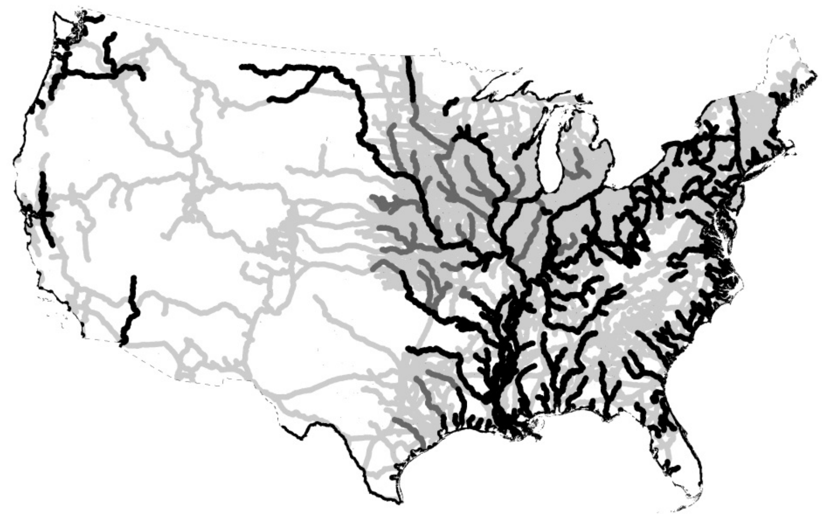
A. 40-Mile Buffers: Waterways (Black) and Railroads (Gray)

B. 40-Mile Buffers: Including Proposed Canals (Dark Gray)

C. 10-Mile Buffers: Waterways (Black) and Railroads (Gray)

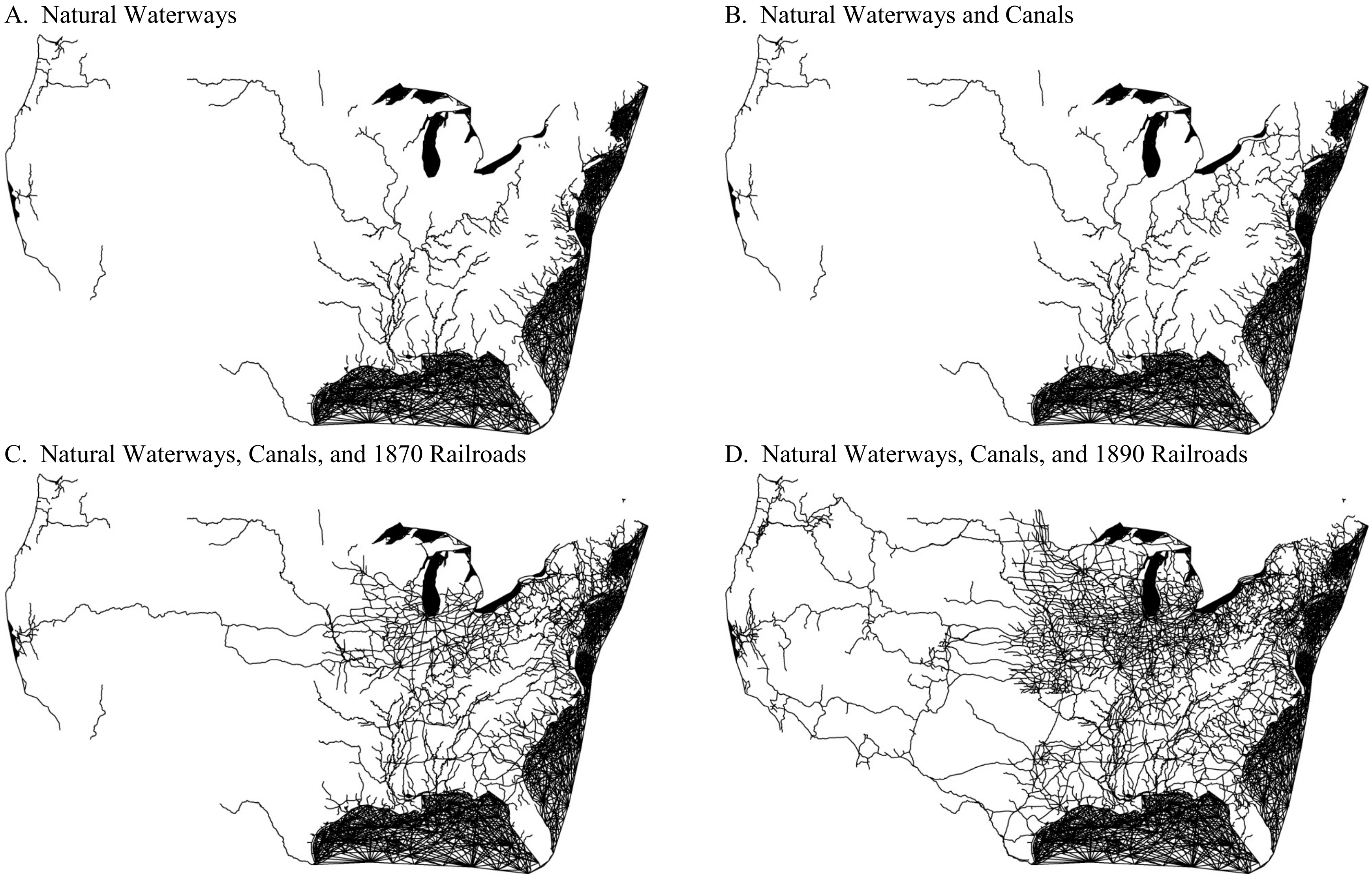


D. 10-Mile Buffers: Including Proposed Canals (Dark Gray)



Notes: In Panel A, areas shaded light gray are within 40 miles of a railroad in 1890 but not within 40 miles of a waterway (shaded black). In Panel B, areas shaded dark gray are further than 40 miles from a waterway but within 40 miles of Fogel's proposed canals. Panels C and D are equivalent for 10-mile buffers.

Figure 2. Constructed Network Database (Partial)



Notes: Panel A shows all natural waterways, including navigable rivers and routes across lakes and oceans. Panel B adds the canal network (as actually constructed in 1890). Panel C adds railroads constructed in 1870, and then Panel D adds railroads constructed between 1870 and 1890.

Table 1. Estimated Elasticity between Market Access and Land Values, By Region

	Log Value of Agricultural Land and Buildings					
	All Counties	By Region:				
	(1)	(2)	(3)	(4)	(5)	(6)
Log Market Access	1.477**	1.895**	1.535**	1.020**	0.819	0.928
	(0.254)	(0.585)	(0.432)	(0.345)	(0.583)	(0.553)
Region	All	Midwest	Plains	South	Northeast	Far West
Number of Counties	2,161	699	192	936	241	93
R-squared	0.587			0.590		

Notes: Column 1 reports estimates from equation (13) in the text. For a balanced panel 2,161 counties in 1870 and 1890, the Log Value of Agricultural Land and Buildings is regressed on Log Market Access, as defined in equation (12). All regressions include county fixed effects, state-by-year fixed effects, and year-specific cubic polynomials in county latitude and longitude. Columns 2 - 6 report estimates from one regression, where the coefficient on Log Market Access is allowed to vary by region. Robust standard errors clustered by state are reported in parentheses: ** denotes statistical significance at the 1% level, and * at the 5% level.

Table 2. Market Access Elasticity: Robustness to Controls for Local Railroads

	Log Value of Agricultural Land and Buildings		Log Market Access	Log Value of Agricultural Land and Buildings	
	(1)	(2)	(3)	(4)	(5)
Log Market Access	1.477**			1.443**	1.455**
	(0.254)			(0.240)	(0.320)
Any Railroad Track		0.359**	0.223**	0.037	0.044
		(0.116)	(0.020)	(0.098)	(0.092)
Railroad Track Length					- 0.032
(Units = 100km)					(0.070)
Number of Counties	2,161	2,161	2,161	2,161	2,161
R-squared	0.587	0.544	0.665	0.587	0.587

Notes: Column 1 reports the baseline specification from Table 1. For the indicated outcome variable, Column 2 and Column 3 report the coefficient on a dummy variable for whether the county contains any railroad track. Column 4 reports the baseline specification from Table 1, controlling for a dummy variable for whether the county contains any railroad track. Column 5 also controls for the length of county railroad track (in units of 100km). All regressions include county fixed effects, state-by-year fixed effects, and year-specific cubic polynomials in county latitude and longitude. Robust standard errors clustered by state are reported in parentheses: ** denotes statistical significance at the 1% level, and * at the 5% level.

Table 3. Market Access Elasticity: Instrumenting with Water Market Access in 1870

	Change in Log Land Value	Change in Log Market Access	Change in Log Land Value	
	OLS	OLS	2SLS	2SLS
	(1)	(2)	(3)	(4)
Change in Log			2.154**	1.271**
Market Access			(0.549)	(0.314)
Log Water Market	-0.668**	-0.310**		
Access in 1870	(0.166)	(0.054)		
Number of Counties	2,161	2,161	2,161	2,161
R-squared	0.561	0.686	0.575	0.772

Notes: For the indicated outcome variable, Column 1 and Column 2 report the impact of Log Water Market Access in 1870. Column 3 reports the estimated impact of a Change in Log Market Access, instrumenting for the Change in Log Market Access with Log Water Market Access in 1870. Column 4 presents the same estimate as in Column 3, but also controlling for Log Land Value in 1870. All regressions (in differences) include state fixed effects and cubic polynomials in county latitude and longitude. Robust standard errors clustered by state are reported in parentheses: ** denotes statistical significance at the 1% level, and * at the 5% level.